

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
12 December 2002 (12.12.2002)

PCT

(10) International Publication Number  
**WO 02/099135 A1**

- (51) International Patent Classification<sup>7</sup>: **C12Q 1/68 //** (81) Designated States (*national*): AE, AG, AL, AM, AT (utility model), AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ (utility model), CZ, DE (utility model), DE, DK (utility model), DK, DM, DZ, EC, EE (utility model), EE, ES, FI (utility model), FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK (utility model), SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.
- (21) International Application Number: PCT/SE02/01093
- (22) International Filing Date: 5 June 2002 (05.06.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
0101999-1 6 June 2001 (06.06.2001) SE  
0103991-6 27 November 2001 (27.11.2001) SE
- (71) Applicant (*for all designated States except US*): CANAG DIAGNOSTICS AB [SE/SE]; Majnabbe Terminal, S-415 55 Göteborg (SE).
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): ÅMAN, Pierre [SE/SE]; Björkvägen 8, S-224 56 Lund (SE). STÅLBERG, Anders [SE/SE]; Ranunkelgatan 7C, lgh 11, S-431 32 Mölndal (SE). KUBISTA, Mikael [SE/SE]; Nedre Solstensvägen 6D, S-435 31 Mölnlycke (SE).
- (74) Agent: STRÖM & GULLIKSSON IP AB; Sjöporten 4, S-417 64 Göteborg (SE).
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— with international search report

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: METHOD TO MEASURE GENE EXPRESSION RATIO OF KEY GENES

(57) Abstract: The invention is a method to determine the amounts, in particular the relative amounts, of nucleic acids in complex biological samples by means of real-time PCR. According to the invention the biological sample is systematically diluted and each dilution is studied by real-time PCR for all genes of interest. From the dependence of the threshold cycle on dilution factor for each of the genes, the PCR efficiencies of the reactions are determined in the particular samples. Determining also the relative sensitivity of the real-time PCR assays compared, the relative amounts of two nucleic acids in complex biological samples are determined with unprecedented accuracy.

WO 02/099135 A1



**TITLE****METHOD TO MEASURE GENE EXPRESSION RATIO OF KEY GENES.****DESCRIPTION**

5 Technical filed

The invention belongs to the category methods for quantification of nucleic acids. Such methods are used to determine the amount of specific genes, gene segments, RNA molecules and other nucleic acids in samples. These methods are primarily used in clinical diagnosis, for example, to test tissue, blood and urine samples, and in food technology,  
10 agriculture and biomedicine.

**Background of the invention**

Methods to measure gene expression go back to the 1970s. The first method was based on measuring reassociation kinetics of complementary strands (Wetmur & Davidson, J. Mol.  
15 Biol., 1, 349, 1968). A radiolabeled single-stranded DNA probe was added and its association with complementary mRNA, when the mRNA was present in molar excess, was measured. These were very difficult experiments, for several reasons: the concentrations of reagents in the hybridization reactions were often so low that the reassociation reaction required many hours – days in some cases – to generate significant amounts of hybrid.  
20 Second, the hydroxyapatite columns routinely used to separate double-stranded and single-stranded nucleic acids were messy to work with, which made the whole procedure tedious. 10 years later Northern hybridization was developed (Alwine, Kemp, & Stark, Proc. Natl. Acad. Sci. U.S.A. 74, 5350, 1977). Here the RNA was immobilized on cellulose and later nitrocellulose paper to which radiolabeled probes were hybridized. The method has several  
25 disadvantages. Its capacity to bind nucleic acids is low and varies according to the size of the RNA. In particular, nucleic acids <400 bases in length are retained inefficiently. Since the RNA is attached to the nitrocellulose by hydrophobic interaction, rather than covalently, it leaches slowly from the matrix during hybridization and washing at high temperatures. Ribonuclease protection assay (Pape, Melchior, & Marotti, Genet. Anal. Tech. Appl. 8, 206,  
30 1991) is 20-100 fold more sensitive than northern hybridization being capable of detecting about  $10^5$  copies of a specific transcript. It can cope with several target mRNAs simultaneously and, because the intensity of the signal is directly proportional to the concentration of target RNA, comparison of the level of expression of the target gene in

different tissues is easily accomplished. A disadvantage is that it works best with antisense probes that are exactly complementary to the target RNA, which is a problem if the experiment generates RNA-RNA hybrids containing mismatched base pairs that are susceptible to cleavage by RNase, for example, when analyzing families of related mRNAs.

5 In 1983 the polymerase chain reaction (PCR) to amplify nucleic acids in an exponential process was invented (US 4,683,202). This opened the possibility to quantify even minute amounts of a nucleic acid in a sample. In traditional PCR the DNA (or RNA after conversion to cDNA) in the sample was amplified first and then detected in a separate step. This made quantification very uncertain, since the reaction usually ran short of some  
10 components giving rise to the same amount of product irrespectively of the amount of starting template.

This problem was solved by inventing real-time PCR (US 6,171,785 ), where fluorescent dyes or fluorescent probes (N. Svanvik, G. Westman, W. Dongyuan & M. Kubista. Anal. Biochem. 281, 26-35, 2000) are included in the reaction to provide for real-time monitoring  
15 of the product formed. The number of amplification cycles required to reach a particular signal threshold level, number of amplification cycles at threshold (CT), is registered. Traditionally the number of template copies in the test sample is estimated by comparing the measured CT value with CT values measured for standard samples containing known amounts of template. This approach is highly accurate when the test sample is of similar  
20 complexity as the standard samples, which usually are dilutions of plasmid or purified DNA template. This relies on the crucial assumption that PCR efficiencies in test and standard samples are the same. If this is not the case a CT-value measured in a test sample will correspond to a different number of cDNA copies than the same CT-value measured in the standard sample. The error introduced by such assumption may be substantial owing to  
25 accumulation effects. For example, 80% efficiency in the test sample and 85% efficiency in the standard sample results in 50% difference in the number of DNA copies after 25 cycles (eq. 1).

$$N_{CT} = N_0 * (1 + E)^{CT}$$

The common method to account for differences in PCR efficiencies between test and  
30 standard samples is to amplify a reference gene, usually a housekeeping gene, in parallel and relating the expression of the studied target gene to the expression of the housekeeping gene. This, of course, relies on the assumption that the expression level of the housekeeping

gene is constant among the samples being compared, which has been questioned (Bustin SA: Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. J Mol Endocrinol 2000, 25:169-193; Suzuki T, Higgins PJ, Crawford DR: Control Selection for RNA Quantitation. BioTechniques 2000, 29:332-337; Schmittgen TD, Zakrajsek B A: Effect of experimental treatment on housekeeping gene expression: validation by real-time, quantitative RT-PCR. J Biochem Biophys Methods 2000, 46:69-81). Further, which is rarely acknowledged, it also assumes that the efficiencies of the two reactions, i.e., the PCR of the target gene and the PCR of the housekeeping gene, are inhibited to the same degree in the standard sample as well as in the test sample (eq. 2):

$$10 \quad \frac{(1 + E_{\text{target gene}}^{\text{test sample}})}{(1 + E_{\text{housekeeping gene}}^{\text{test sample}})} = \frac{(1 + E_{\text{target gene}}^{\text{standard sample}})}{(1 + E_{\text{housekeeping gene}}^{\text{standard sample}})}$$

The validity of this critical assumption has not been tested, because there has been no method to determine the PCR efficiencies of individual reactions in samples.

One object of the present invention is to overcome the limitations discussed above with traditional methods to determine gene expression and also the limitations of the present real-time PCR approach to quantify the relative amounts of two nucleic acids in a biological sample.

Another object of the present invention is to diagnose a disease, such as cancers and in particular lymphomas, with very high sensitivity by measuring the ratio of expression of key genes.

Still another object of the present invention is to diagnose a disease with technology that requires very little material as obtained, for example, with fine needle aspiration biopsy. Still another object of the present invention is to make diagnosis rapid and more cost efficient.

#### Description of figures

**Figure 1. Controlled dilution of test sample.** The test sample is diluted 64 times in three steps á four times.

**Figure 2. Inter and intra assays.** Top left: IgL $\kappa$  intra assay; top right: IgL $\lambda$  intra assay; bottom left: IgL $\kappa$  inter assay; bottom right: IgL $\lambda$  inter assay.

**Figure 3. Variations in inter and intra assays.** Variations in CT-values for the IgL $\kappa$  and IgL $\lambda$  reactions in eight repeated measurements run either in parallel (intra-assay) or separately (inter-assay) of sample BR0.

**Figure 4. PCR efficiencies of the IgL $\kappa$  (A) and IgL $\lambda$  (B) assays.** The lines are normalized at maximum template concentration. PCR efficiencies are obtained from the slopes of the fitted lines as  $E = 10^{-(\text{slope})^{-1}} - 1$ . The outlier, sample BR17, is indicated with dotted line (···). Purified template is shown with dashed line (- - -). For all lines  $R^2 > 0.99$ .

**Figure 5. IgL $\kappa$  and IgL $\lambda$  PCR efficiencies in lymphoma samples.** PCR efficiencies of the IgL $\kappa$  and IgL $\lambda$  reactions determined by the invented approach in seven test samples and of purified template. The calculated relative sensitivity,  $K_{RS}$ , in the negative samples is also shown.

**Figure 6. Classification of lymphoma samples.** Patient samples shown in a CT $\kappa$  vs. CT $\lambda$  plot. Each symbol represents one sample and is depicted at its CT $\kappa$  and CT $\lambda$  values. The opposite axes indicate the number of cDNA copies for purified template. The straight solid line represents (CT $\kappa$ , CT $\lambda$ ) values expected for negative samples calculated assuming 85.4% and 79.3% PCR efficiencies for the IgL $\kappa$  and the IgL $\lambda$  reactions, respectively. The dotted lines (···) indicate an interval within which negative samples should be found with at least 95 % probability. B-cell lymphomas are shown with ■, diffuse large B-cell lymphoma with \* and negative samples with ●. Open symbols indicate corrected CT-values of samples for which specific PCR efficiencies were determined.

**Figure 7. Comparison of classification by various methods of NHL samples.** Classification of patient samples by the invented real-time PCR method compared with traditional R.E.A.L. classification, classification by IHC clonality and by flow cytometry. Positive B cell lymphoma samples are shown in bold. The more rapid and for the patient less inconvenient invented real-time PCR method does in all cases agree with the traditional methods.

**Figure 8. Determination of PCR efficiencies for bcr-abl and GAPDH using probes.** CT values measured for the bcr-abl and GAPDH reactions using Taqman probe real-time PCR assays in a patient sample systematically diluted in steps of two. The CT v.s. log(dilution) plots have different slopes evidencing that the two reactions are inhibited to different degrees in the sample. The ratio between bcr-abl and GAPDH cDNA are calculated taking the CR efficiencies into account.

**Figure 9. PCR efficiencies of bcr-abl and GAPDH reactions in patient samples.** Table showing the PCR efficiencies of the bcr-abl and GAPDH reactions measured using Taqman probe real-time PCR assays in five patient samples determined by the invented method. In all samples was the GAPDH reaction inhibited to a higher degree. The degree of inhibition of both reactions also vary substantially among the samples evidencing the importance of the present invention.

**Figure 10. Determination of bcr-abl cDNA using dye.** Real-time PCR amplification curves of a SYBRGreen assay of bcr-abl cDNA. Top left shows plot of CT v.s. log(starting concentration) and top right shows melting curves distinguishing template specific products from primer dimers.

**Figure 11. Determination of GAPDH cDNA using dye.** Real-time PCR amplification curves of a SYBRGreen assay of GAPDH cDNA. Top left shows plot of CT versus log(starting concentration) and top right shows melting curves distinguishing template specific products from primer dimers.

#### **Summary of the present invention**

The present invention is a method to determine the relative amounts of two nucleic acids, in particular two cDNAs, in complex biological samples by real-time PCR. It is based on determining the threshold cycles (CT) of the PCR:s of a dilution series of the test sample, and from the dependences of CT on the logarithm of the dilution factor determine the PCR efficiencies of the two reactions in the particular sample.

With the here invented method it is possible to determine PCR efficiency in biological test samples.

With the here invented method it is possible to determine the ratio of two nucleic acids in biological test samples with unprecedented accuracy by taking into account the sample  
5 specific inhibition.

With the here invented method it is possible to determine the ratio of two cDNA and thereby indirectly of the corresponding mRNAs and, hence, the relative expression of two genes.

10 With the here invented method it is possible to determine the ratio of the expression of IgL $\kappa$  and IgL $\lambda$  genes thereby detecting clonality of B cells and classifying lymphoma.

The fundamental inventive idea is that the sample itself is used as a standard reference by using a dilution or a concentrate thereof as comparative standard.

15 Detailed description of the invention and its preferred embodiments

As indicated by the title, the present invention is a procedure to determine the ratio of two nucleic acids, in particular of two cDNAs and hence mRNAs, in complex biological samples by quantitative real-time PCR. As already mentioned the state-of-the art approach expresses the amount of a nucleic acid in a sample relative to the amount of another nucleic  
20 acid. This is the typical case both when measuring viral loads as well as gene expression levels. Typically the expression of the gene of interest is expressed relative to the expression of a house keeping gene, which is a gene assumed to be expressed to the same degree under essentially all conditions. This relative expression of two genes relies on the assumption that the two PCR:s are inhibited to the same degree in the standard sample as well as in the test  
25 sample (eq. 0). So far it has not been possible to test this assumption, because there has been no way to determine PCR efficiencies in individual samples. This is made possible with the invention described here.

Although one might be inclined to think that inhibitory components that may be present in  
30 biological samples should have the same effect on all PCR:s, it may not necessarily be so. The degree of inhibition may depend on features that are particular for the different PCR systems, such as the length and sequence of template, template tertiary structure, lengths and sequences of primers etc. Inhibition may also be indirect through competition for critical



elements such as ions and dNTPs. If two PCR systems have optimum efficiencies at different concentrations of  $Mg^{2+}$ , dNTP, primers and dye/probe elements in biological samples that interact with these PCR components may interfere with the reactions to different degrees.

- 5 The invented approach is based on taking the test sample and performing a controlled dilution, for example, as illustrated in Figure 1, in four steps a four times. By amplifying the nucleic acid in each of these dilutions and comparing the number of cycles required to reach threshold (CT) with the dilution factors, the efficiency of the PCR in that particular sample can be determined. For example, if the reaction proceeds with 100 % efficiency, 4 times
- 10 dilution should increase the CT exactly by 2, 16 times dilution by four and 64 times dilution by 8. From a plot of CT vs. log(dilution factor) the efficiency of the reaction in that particular sample is determined. When comparing the expression of two genes in a biological test sample, the test sample (after cDNA synthesis) is serially diluted and the amounts of both cDNAs are determined in each dilution, from which the PCR efficiencies
- 15 of both reactions in that particular sample are determined.

A mathematical model is developed to determine the ratio of the expression levels of two genes by real-time PCR. The model is general and applied here on the IgL $\kappa$  and IgL $\lambda$  genes. In the following equations the following meanings are due:

- 20  $N_{0A}$  means the number of units,  $N_A$ , at the time 0 of cDNA of type A  
 $N_{0B}$  means the number of units,  $N_B$ , at the time 0 of cDNA of type B  
 $K_{RS}$  means the constant based on relative sensitivity for optical detection  
 $E_A$  means PCR efficiency of sample A  
 $E_B$  means PCR efficiency of sample B
- 25  $[E_A]$  means PCR mean efficiency determined on a larger number of samples of A  
 $[E_B]$  means PCR mean efficiency determined on a larger number of samples of B  
 $CT_A$  means the number of cycles of amplifications in reaction of sample A to reach threshold value.  
 $CT_B$  means the number of cycles of amplifications in reaction of sample B to reach
- 30 threshold value.

The basic equation describing real-time PCR amplification in exponential phase is (eq. 3):

$$N_{CT} = N_0 * (1 + E)^{CT}$$

$N_0$  is the number of cDNA molecules,  $E$  is the PCR efficiency ( $E = 1$  corresponds to 100% efficiency and is expressed in percentage throughout),  $CT$  is the threshold cycle and  $N_{CT}$  is the number of template copies present after  $CT$  PCR cycles.  $E$  is assumed to be independent  
 5 of  $N$  in the particular amplification range. It is determined by performing a dilution series of mRNA or cDNA standard and is calculated from the slope in a  $CT$  vs.  $\log N_0$  plot (eq. 4):

$$E = 10^{-(\text{slope})^{-1}} - 1$$

The fluorescence increase, i.e., the fluorescence signal after subtraction of background, at  
 10 threshold is proportional to the amount of target DNA (eq. 5):

$$I = k * N_{CT}$$

$k$  is a system and instrument constant and  $N_{CT}$  is the number of target DNA molecules present at threshold. The relative expression of the  $IgL\kappa$  and  $IgL\lambda$  genes is obtained as (eq. 6, eq. 7, eq. 8, and eq. 9)

15

$$N_{CT_{IgL\kappa}} = N_{0_{IgL\kappa}} * (1 + E_{IgL\kappa})^{CT_{IgL\kappa}}$$

$$I_{IgL\kappa} = k_{IgL\kappa} * N_{CT_{IgL\kappa}}$$

$$N_{CT_{IgL\lambda}} = N_{0_{IgL\lambda}} * (1 + E_{IgL\lambda})^{CT_{IgL\lambda}}$$

$$I_{IgL\lambda} = k_{IgL\lambda} * N_{CT_{IgL\lambda}}$$

20

At threshold  $I_{IgL\kappa} = I_{IgL\lambda}$ . Equating eq. 5 with eq. 7 and rearranging we obtain (eq. 10):

$$K_{RS} = \frac{k_{IgL\lambda}}{k_{IgL\kappa}} = \frac{N_{CT_{IgL\kappa}}}{N_{CT_{IgL\lambda}}}$$

where the relative sensitivity,  $K_{RS}$  reflects the difference in probes' fluorescence and binding efficiencies in the two assays. Inserting eq. 4 and 6 and rearranging we get (eq. 11):

25

$$\frac{N_{0_{IgL\kappa}}}{N_{0_{IgL\lambda}}} = K_{RS} * \frac{(1 + E_{IgL\lambda})^{CT_{IgL\lambda}}}{(1 + E_{IgL\kappa})^{CT_{IgL\kappa}}}$$

This is the central equation to calculate the ratio between the numbers of copies of two cDNA molecules.  $CT_{IgL\kappa}$  and  $CT_{IgL\lambda}$  are the  $CT$  values obtained from the PCR

amplifications of the IgL $\kappa$  and IgL $\lambda$  cDNAs,  $E_{\text{IgL}\kappa}$  and  $E_{\text{IgL}\lambda}$  are the efficiencies of the two PCR equations determined as slopes in plots of CT vs.  $\log N_0$  in the serial dilutions of the samples, and  $K_{\text{RS}}$  is the relative sensitivity constant of the two PCR assays determined using test samples with known cDNA concentrations.

- 5 The fractions of IgL $\kappa$  and IgL $\lambda$  mRNA expressed as percentage are finally calculated as (eq. 12, and eq. 13):

$$\text{IgL}\kappa = 100 * \frac{K_{\text{RS}} * \frac{(1 + E_{\text{IgL}\lambda})^{CT_{\text{IgL}\lambda}}}{(1 + E_{\text{IgL}\kappa})^{CT_{\text{IgL}\kappa}}}}{1 + K_{\text{RS}} * \frac{(1 + E_{\text{IgL}\lambda})^{CT_{\text{IgL}\lambda}}}{(1 + E_{\text{IgL}\kappa})^{CT_{\text{IgL}\kappa}}}}$$

$$\text{IgL}\lambda = 100 * \frac{1}{1 + K_{\text{RS}} * \frac{(1 + E_{\text{IgL}\lambda})^{CT_{\text{IgL}\lambda}}}{(1 + E_{\text{IgL}\kappa})^{CT_{\text{IgL}\kappa}}}}$$

- 10 To determine PCR efficiencies in a biological sample by studying the effect of dilution on CT, the experimental variation in CT due to experimental uncertainty and variation in PCR efficiency owing to added components must be small compared to that caused by dilution. We established this to be the case by determining the experimental reproducibility using a typical patient sample that was analyzed for expression of the immunoglobulin kappa and
- 15 lambda light chain in example 1. The PCR efficiencies in the biological samples are according to this invention determined by first converting the mRNA to cDNA and then serially diluting the sample determining the CT values of both reactions after each dilution. A single dilution is sufficient to estimate PCR efficiency, but the more dilutions made the higher is the accuracy. However, too extensive dilutions should be avoided, because if the
- 20 number of molecules gets too few stochastic errors may be introduced (Vogelstein B, Kinzler KW: Digital PCR. Proc Natl Acad Sci USA 1999, 96: 9236-9241; Peccoud J, Jacob C: Theoretical uncertainty of measurements using quantitative polymerase chain reaction. Biophys J 1996, 71: 101-108). In example 2 we diluted 64 times in three steps of four times, which changed CT sufficiently to make experimental errors negligible. We also used
- 25 samples that contained at least 6500 molecules of each cDNA, corresponding to at least 100 cDNAs of each in the most diluted sample.

#### Application in cancer diagnostics

Cancer is tissue that grows uncontrolled. The cancer cells have lost control of their cell division mechanism and divide indefinitely. All cancer cells originate in a single cell that has gone awry. In this cell genes that should be silent are active, and it often also loses  
5 ability to express growth controlling genes or expresses aberrant or foreign genes. Since all cancer cells originate from the same cell they share genetic signature, which can be used to detect and diagnose the cancer.

Particular kinds of cancer are lymphomas, which are cancers of the lymphatic system. Like  
10 other cancers lymphomas occur when cells divide too much and too fast. Growth control is lost, and the lymphatic cells may overcrowd, invade, and destroy lymphoid tissues and metastasize (spread) to other organs. There are two general types of lymphomas: "Hodgkin's Disease" (named after Dr. Thomas Hodgkin, who first recognized it in 1832) and non-Hodgkin's lymphoma (NHLs). Non-Hodgkin's Lymphomas caused by malignant (cancerous)  
15 B-cell lymphocytes represent a large subset (about 85% in the US) of the known types of lymphoma (the other two subsets being T-cell lymphomas and lymphomas where the cell type is unknown).

The traditional way to diagnose lymphoma is to take a surgical biopsy and test it by  
20 immunocytochemistry, flow cytometry and cytogenic studies. These tests rely on cell-specific antibodies. As alternative a fine needle aspiration (FNA) biopsy could be taken. This uses a very thin, hollow needle that is attached to a syringe. The needle is inserted into the swollen lump. It is then pushed back and forth to free some cells, which are aspirated (drawn up) into the syringe. FNA can distinguish noncancerous conditions, like infections,  
25 from NHLs or other cancers. FNA also is useful for staging, or determining the extent, of disease, and for monitoring recurrence, or return of cancer. But, because of small sample sizes and lack of information about lymph node structure, FNA often is inadequate for the initial diagnosis of NHL using current immunologic methods. A great improvement would be a more sensitive method than those based on immunochemistry, for which material from  
30 FNA would be sufficient.

B-lymphocytes produce immunoglobulins having a heavy chain and either a kappa (IgL $\kappa$ ) or a lambda (IgL $\lambda$ ) light chain. Each B-lymphocyte decides early in its development which

light chain to produce. In healthy humans about sixty per cent of the B-cells produce kappa chains and the rest produce lambda chains. Normal lymphoid tissues therefore contain a mixture of B-cells with a  $IgL\kappa : IgL\lambda$  ratio of about 60:40 (Levy R, Warnke R, Dorfman RF, Haimovich J: The monoclonality of human B-cell lymphomas. J Exp Med 1977, 154:1014-1028; Barandun S, Morell A, Skvaril F, Oberdorfer A: Deficiency of kappa- or lambda-type immunoglobulins. Blood 1976, 47:79-89). Lymphomas, like all malignant tumors, are clonal and arise from one transformed cell. Lymphoma tissues are dominated by the tumor cells and consequently the  $IgL\kappa : IgL\lambda$  ratio is changed. Kappa producing tumors result in a higher  $IgL\kappa : IgL\lambda$  ratio, while lambda producing tumors result in a lower ratio. Assuming that the translation efficiency and stability of the  $IgL\kappa$  and  $IgL\lambda$  mRNAs are similar, clonality may be detected by measuring the  $IgL\kappa : IgL\lambda$  expression ratio. In Example 3 we show how patient samples can be classified as NHL positive and NHL negative from the determined  $IgL\kappa : IgL\lambda$  expression ratio by the method invented here. The excellent accuracy is impressive in view of the very little amount of material needed for analysis. The 1000 to 100000 representative cells typically obtained in a fine needle aspiration biopsy are sufficient for at least 50 tests by the real-time PCR assay and detection of possible B-cell monoclonality in the specimen by the invention presented here.

Another possible application of the method invented here is to detect T cell clonality. Here instead markers will be variants of the T cell receptors

Still another application of the method invented here is to monitor progress of disease. Some cancers are caused by expression of unnatural proteins, such as the bcr-abl fusion protein in Chronic Myelogenous/Myeloid Leukemia (CML) patients. It is important to quantify the amount of bcr-abl fusion transcript for diagnosis, and it is even more important to monitor disease progress. Imatinib mesylate (Gleevec® also known STI571) is a molecule in clinical trials for treatment of CML patients and to optimize treatment it is desired to know how patients respond to the drug, which is measured as changes in bcr-abl expression. Since drug treatment may affect overall gene expression, the expression of bcr-abl is usually determined relative to a house-keeping gene such as GAPDH. In Example 4 we show that bcr-abl and GAPDH PCR efficiencies are inhibited to different degree in CML patient sample and, hence, the importance of taking this into account when determining expression ratios and effect of drug treatment.

Indeed any diagnosis based on determining gene expression levels are possible applications of the method invented here. It is not limited to determining the ratio of expression of two genes; some diseases may be characterized by a particular expression pattern of three or  
5 even more genes.

Another possible application of the method invented here is to measure the relative amount of various splicing variants of a gene, which may be of interest in diagnosis or prognosis. The PCR efficiencies of the various splicing variants, which in general differ in both lengths  
10 and sequence, may vary, and correction may be important to obtain an accurate measure. Another possible application of the method invented here is to measure the relative activity of alternative promoters of genes. These are also likely to be amplified with different efficiencies that should be taken into account for proper diagnosis and prognosis.

## 15 Examples

### Example 1. Experimental reproducibility.

Surgical lymph node biopsies from previously untreated patients were transported from the operation theatre in ice water chilled boxes and handled in the laboratory within 30 minutes. Material for the study was rapidly frozen in dry ice /isopentane and stored at  $-70^{\circ}\text{C}$ .  
20 Parts of the tissues were fixed in formalin and used for routine diagnostic analysis. Diagnosis was reached by a combination of microscopic evaluation of histology, immunostaining of several markers including the kappa and lambda chains (IHC) and in some cases flow cytometry. The samples were classified as lymphadenitis or malignant lymphoma according to the R.E.A.L.-terminology (Harris NH, Jaffe ES, Stein H, Banks PM, Chan JK, Cleary ML, Delsol G, De Wolf-Petters C, Falini B, Gatter KC: A proposal  
25 from the International Lymphoma Study Group. Blood 1994, 84:1361-1392).

RNA was extracted using the Fast Prep System (FastRNA Green, Qbiogene). Ten  $\mu\text{g}$  of total RNA was mixed with 2  $\mu\text{g}$  of pdT oligomers (Pharmacia) and incubated at  $65^{\circ}\text{C}$  for 5  
30 minutes. First strand cDNA synthesis was then performed by adding 0.05 M tris-HCl, pH 8.3, 0.075 M KCl, 3 mM  $\text{MgCl}_2$ , 0.01 M DTT, 10 U/ml M-MLV reverse transcriptase (Life Technologies), 0.05 U/ml RNA guard (Life Technologies) and 10 mM of each deoxyribonucleotide to a final volume of 20 ml and incubating the samples at  $37^{\circ}\text{C}$  for one

hour. The reaction was terminated by incubation at 65°C for 5 minutes and samples were stored at -80°C.

Two homopyrimidine light-up probes, H-CCTTTTCCCC-NH<sub>2</sub> (IgLκLUP) and  
5 CCTCCTCTCT-NH<sub>2</sub> (IgLλLUP), directed against PCR amplification products of the  
constant regions in the human immunoglobulin kappa (IgLκ) and lambda (IgLλ) light-chains  
respectively, were designed. Both probes are homopyrimidine sequences, which are known  
to exhibit very large signal enhancement upon target binding (Svanvik N, Nygren J,  
Westman G, Kubista M: Free-probe fluorescence of light-up probes. *J Am Chem Soc* 2001,  
10 123:803-809). Both probes had the thiazole orange derivate, N-carboxypentyl-4- [(3'-  
methyl-1', 3'-benzothiazol-2'-yl) methylenyl] quinolinium iodide (TO-N-5-COOH), as  
label. They were synthesized by solid phase synthesis and purified twice by reverse phase  
HPLC as described (Svanvik N, Westman G, Wang D, Kubista M: Light-up probes: thiazole  
orange-conjugated peptide nucleic acid for detection of target nucleic acid in homogeneous  
15 solution. *Anal Biochem* 2000, 281:26-35). Probe concentrations were determined  
spectroscopically assuming molar absorptivities at 260 nm of 83,100 M<sup>-1</sup>cm<sup>-1</sup> for IgLκLUP  
and 81,100 M<sup>-1</sup>cm<sup>-1</sup> for IgLλLUP.<sup>7</sup> The probes were designed to have melting temperatures  
(T<sub>m</sub>) of 65-70°C, which is in between the annealing (T<sub>annealing</sub> = 55°C) and elongation  
(T<sub>elongation</sub> = 74°C) temperatures of the PCR:s.

20 PCR products were purified by QIAquick<sup>TM</sup> PCR purification kit (Qiagen) and their  
concentrations were determined spectroscopically assuming molar absorptivity at 260 nm of  
13,200 M<sup>-1</sup>cm<sup>-1</sup> per base pair. Primer (Medprobe Inc) concentrations were estimated  
assuming  $\epsilon_{260}/10^3 = 12.0n_G + 7.1n_C + 15.2n_A + 8.4n_T$  M<sup>-1</sup>cm<sup>-1</sup>, where n<sub>x</sub> is the total number  
25 of base x (Current Protocols in Molecular Biology. Edited by Ausubel FM, Brent R,  
Kingstone R, Moore DD, Seidman JG, Smith JA, Struhl K. John Wiley & Sons, Inc.  
Canada, 2000, pp. A.3D.2)

PCR systems were designed for a 231bp fragment of the human IgLκ (GenBank accession  
30 number AK024974) and a 223bp fragment of the human IgLλ (GenBank accession number  
X51755) comprising the IgLκLUP and IgLλLUP target sequences, respectively. Reaction  
conditions were optimized as described elsewhere (Kubista M, Ståhlberg A, Bar T: Light-up  
probe based real-time Q-PCR. *Proceedings of SPIE, in Genomics and Proteomics*

Technologies, Raghavachari R, Tan W, Editors. Proceedings of SPIE 2001, 4264:53-58).  
IgL $\kappa$  and IgL $\lambda$  PCR:s both contained 75 mM Tris (pH 8.8), 20 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.1% Tween  
20, 1 U of JumpStart™ Taq DNA polymerase (with antibody) (Sigma-Aldrich) and 200  
ng/ $\mu$ L of BSA. Specific components for the IgL $\kappa$  PCR were 5mM MgCl<sub>2</sub>, 0.2mM  
5 deoxyribonucleotides (Sigma-Aldrich), 800nM of each primer (MedProbe) and 800nM  
IgL $\kappa$ LUP, and for the IgL $\lambda$  PCR 3.5 mM MgCl<sub>2</sub>, 0.4mM deoxyribonucleotides, 600 nM of  
each primer and 600 nM IgL $\lambda$ LUP. Primer sequences were for IgL $\kappa$  5'-TGA GCA AAG  
CAG ACT ACG AGA-3' (forward) (SEQ. ID. NO.1) and 5'-GGG GTG AGG TGA AAG  
ATG AG-3' (reverse) (SEQ. ID. NO. 2), and for IgL $\lambda$  5'- GAG CCT GAC GCC TGA G -  
10 3'(forward) (SEQ. ID. NO. 3) and 5'- ATT GAG GGT TTA TTG AGT GCA G-3' (reverse)  
(SEQ. ID. NO. 4).

Real-time PCR was measured in a LightCycler (Roche Diagnostics) using the thermocycler  
program: 3 min pre-incubation at 95°C followed by 50 cycles for 0 s at 95°C, 10 s at 55°C  
15 and 11 s at 74°C. Fluorescence was monitored at the end of the annealing phase using 470  
nm excitation and 530 nm emission (the LightCycler F1 channel). All amplification curves  
were baseline adjusted by subtracting the arithmetic average of the five lowest fluorescence  
read-out values in each sample (arithmetic baseline adjustment in the LightCycler software).  
The threshold was set to a value of 1.00, which was significantly above background noise,  
20 and the number of cycles required to reach this level, CT, was determined (Higuchi R,  
Fockler C, Dollinger G, Watson R: Kinetic PCR analysis: real-time monitoring of DNA  
amplification reactions. Biotechnology (N Y) 1993, 11:1026-1030).

To classify a sample as either lymphoma negative with 60:40 IgL $\kappa$  : IgL $\lambda$  expression ratio or  
positive with a deviating expression ratio, we must know with what accuracy CT can be  
25 determined. We therefore designed experiments to measure the variation in CT due to  
experimental error and biological variability. First we studied the reproducibility of the PCR  
by splitting a sample into aliquots that were analyzed in parallel runs (intra-assay). We then  
also included variation due to sample handling by analyzing the same sample in independent  
runs (inter-assay). To minimize variation in template concentration between the two assays  
30 being compared a master mix containing template and all common PCR components was  
prepared and split into two aliquots to which the unique components for the IgL $\kappa$  and the  
IgL $\lambda$  reactions were added. Each experiment was performed 8 times using patient sample  
BR0 (figure 2).



In most reports PCR reproducibility is expressed as standard deviation in CT. The variance,  $SD^2$ , is (eq. 14)

$$SD^2 = \frac{\sum_{i=1}^n (CT_i - \langle CT \rangle)^2}{n-1}$$

- 5 where  $\langle CT \rangle$  is the average of the measured CT and standard deviation, SD, is the square root of the variance. However, since we are interested in determining the amount of cDNAs in the sample, the standard deviation of  $(1+E)^{-CT}$ , which is proportional to the number of cDNA molecules (eq. 1, and eq. 15):

$$N_0 = N_{CT} * (1+E)^{-CT}$$

- 10 is more relevant. The variance in  $(1+E)^{-CT}$  is (eq. 16)

$$SD^2 = \frac{\sum_{i=1}^n ((1+E)^{-CT_i} - \langle (1+E)^{-CT} \rangle)^2}{n-1}$$

where  $\langle (1+E)^{-CT} \rangle$  is the average of  $(1+E)^{-CT}$ . To obtain the relative uncertainty in the number of cDNA molecules, we normalize the standard deviation with the average value to obtain the coefficient of variation, CV, which we express in percent (eq. 17):

15  $CV = 100 \times SD / \langle (1+E)^{-CT} \rangle$

- CV is the uncertainty in the determination of the number of cDNA molecules in the sample due to experimental factors. In the intra-assay, which reflects the reproducibility of the PCR, the coefficient of variation was 3.0% for the IgL $\kappa$  reaction and 4.9 % for the IgL $\lambda$  reaction (Figure 3). For the inter-assay, where also experimental errors contribute, the coefficients of variation were only slightly larger; 8.1% for the IgL $\kappa$  reaction and 5.0% for the IgL $\lambda$  reaction. Although it is not possible to calculate a coefficient of variation for the ratio of the two cDNAs we can estimate how much the IgL $\kappa$  : IgL $\lambda$  expression ratio in a negative sample could deviate from 60:40 due to experimental uncertainty in a bad case. Suppose the number of IgL $\kappa$  cDNA is overestimated due experimental error by one standard deviation and the number of IgL $\lambda$  cDNA is underestimated also by one standard deviation the measured ratio would be  $(60/40) \times (1+0.081)/(1-0.050) = 1.70 = 63/37$ . If instead the amount of IgL $\kappa$  cDNA is underestimated and that of IgL $\lambda$  cDNA is overestimated then the
- 25

measured ratio would be  $(60/40) \times (1-0.081)/(1+0.050) = 1.31 = 56/44$ . Hence, due to experimental uncertainty and variation in PCR efficiency owing to added components we expect negative samples to display an  $IgL\kappa : IgL\lambda$  expression ratio of  $56:44 < N_{0_{IgL\kappa}} : N_{0_{IgL\lambda}} < 63:37$ .

5

#### Example 2. Determination of $IgL\kappa$ and $IgL\lambda$ PCR efficiencies in patient samples

PCR efficiencies in seven patient samples were determined by diluting the test samples in steps and measuring CT value at each dilution. From these data intrinsic standard curves were constructed from which the PCR efficiencies are determined (figure 3). We chose to dilute the samples 64 times, in three steps of 4 times. The dilutions were performed in duplicates and the CT values were measured for both the  $IgL\kappa$  and  $IgL\lambda$  reactions determining the efficiencies of the two assays separately. Seven patient samples, four negative and three positive, were characterized this way, as well as purified template that should not contain any inhibitors.

15

The PCR efficiencies obtained when amplifying purified template were  $E_{IgL\kappa} = 94.7\%$  and  $E_{IgL\lambda} = 93.2\%$  signifying that both reactions proceed with very high efficiencies as expected for optimised PCR assays. Six of the patient samples exhibited efficiencies that were about 10% lower; the  $IgL\lambda$  PCR efficiency was  $75.2\% < E_{IgL\lambda} < 85.8\%$  with mean  $\langle E_{IgL\lambda} \rangle = 79.3\%$  and the  $IgL\kappa$  efficiency was  $79.4\% < E_{IgL\kappa} < 90.4\%$  with mean  $\langle E_{IgL\kappa} \rangle = 85.4\%$  (Table 2). The seventh sample, BR17, exhibited normal  $IgL\kappa$  efficiency (83.0 %), while the  $IgL\lambda$  efficiency was only 58.9 %. The reason for the extremely low efficiency of the  $IgL\lambda$  reaction in this sample is unclear. It was considered outlier and was not included in the calculation of average efficiencies.

25

When comparing the yields of two reactions the efficiency ratio (eq. 18)

$$X_{ER} = \frac{(1 + E_{IgL\kappa})}{(1 + E_{IgL\lambda})}$$

is the relevant parameter (see eq. 9). For the six samples  $1.01 < X_{ER} < 1.065$  with  $\langle X_{ER} \rangle = 1.034$  (Figure 5). Hence, after some 20 amplification cycles, which was typically required to reach threshold with the patient samples (Figure 2), twice  $(1.034^{20} = 2)$  as many kappa DNA

30

molecules have been formed compared to lambda DNA due to the difference in PCR efficiencies.

- Finally, to relate the measured CT-values of the two real-time PCR reactions to the ratio  
 5 between the numbers of corresponding cDNA molecules, we must also determine the relative sensitivity,  $K_{RS}$ , of the two probing systems (eq. 8, and eq. 19).

$$K_{RS} = \frac{N_{0_{IgL\kappa}} * (1 + E_{IgL\kappa})^{CT_{IgL\kappa}}}{N_{0_{IgL\lambda}} * (1 + E_{IgL\lambda})^{CT_{IgL\lambda}}}$$

- was calculated from the CT values ( $CT_{IgL\kappa}$ ,  $CT_{IgL\lambda}$ ) and PCR efficiencies ( $E_{IgL\kappa}$ ,  $E_{IgL\lambda}$ )  
 determined for the four negative samples (table 2) assuming 60:40 IgL $\kappa$  : IgL $\lambda$  expression  
 10 ratio. This gave  $1.41 \leq K_{RS} \leq 1.84$  with mean  $\langle K_{RS} \rangle = 1.52$  (Figure 5). As alternative  $K_{RS}$  was determined using purified template, which concentration was determined spectroscopically, that was diluted and amplified. Hence, the probing of IgL $\kappa$  DNA is about 50% more sensitive than probing of IgL $\lambda$  DNA using the probes and conditions here.

### 15 Example 3. Classification of NHL lymphoma patient samples

- A total of 20 patient samples were analyzed for B-cell lymphoma by the Q-PCR assay. All  
 samples were run in duplicates including negative controls. The data plotted in Figure 6 and  
 summarized in Figure 7. In the plot each symbol represents one sample and is positioned on  
 the coordinates  $CT_{IgL\kappa}$ ,  $CT_{IgL\lambda}$ . The corresponding number of cDNA molecules of purified  
 20 template, calculated assuming  $E_{IgL\kappa} = 94.7\%$  and  $E_{IgL\lambda} = 93.2\%$ , is indicated in logarithmic scale on the opposite axes. Samples considered negative by IHC analysis are shown as circles and positive samples are shown as squares.

- Negative samples with IgL $\kappa$  : IgL $\lambda$  gene expression ratio of 60:40 are expected to lie on a  
 25 straight line. Rewriting equation (9) gives (eq. 20):

$$N_{0_{IgL\kappa}} * (1 + E_{IgL\kappa})^{CT_{IgL\kappa}} = K_{RS} * N_{0_{IgL\lambda}} * (1 + E_{IgL\lambda})^{CT_{IgL\lambda}}$$

converting it to logarithmic form (eq. 21):

$$CT_{IgL\kappa} * \log(1 + E_{IgL\kappa}) = \log(K_{RS} * \frac{N_{0_{IgL\lambda}}}{N_{0_{IgL\kappa}}}) + CT_{IgL\lambda} * \log(1 + E_{IgL\lambda})$$

and rearranging, we obtain (eq. 22):

$$CT_{IgL\kappa} = \frac{\log(1 + E_{IgL\lambda})}{\log(1 + E_{IgL\kappa})} * CT_{IgL\lambda} + \frac{\log(K_{RS} * \frac{N_{0IgL\lambda}}{N_{0IgL\kappa}})}{\log(1 + E_{IgL\kappa})} = CT_{IgL\kappa} = k * CT_{IgL\lambda} + l$$

This describes a linear relation between  $CT_{IgL\kappa}$  and  $CT_{IgL\lambda}$  with slope  $k$  and intercept  $l$ . Inserting  $\langle E_{IgL\kappa} \rangle = 0.854$ ,  $\langle E_{IgL\lambda} \rangle = 0.793$  and  $\langle K_{RS} \rangle = 1.52$ , which are the average values determined for the six samples above (Figure 5), and  $N_{0IgL\kappa} / N_{0IgL\lambda} = 60:40 = 1.5$ , we obtain  $k = 0.946$  and  $l = 0.021$ . Note that the relative sensitivity,  $K_{RS}$ , was calculated from measurements on negative samples assuming 60:40 expression ratio (eq. 17). This cancels the  $N_{0IgL\kappa} / N_{0IgL\lambda}$  ratio in the nominator in the second term. Hence, the calculated slope and intercept of the relation between  $CT_{IgL\kappa}$  and  $CT_{IgL\lambda}$  for negative samples is independent of the assumption of a particular  $IgL\kappa : IgL\lambda$  expression ratio. A line with  $k = 0.946$  and  $l = 0.021$  is drawn in figure 6.

Some negative samples are slightly off the line representing 60:40 expression (Figure 5). This may be due to variations in PCR efficiencies among the samples. Such variations will cause an error in the estimation of the number of cDNA molecules from the measured CT-values when mean PCR efficiencies are assumed. If the efficiencies of the two PCR assays in a sample deviate from the mean values to about the same degrees, the measured CT-values will still correctly reflect the expression ratio and negative samples will fall on the 60:40 line, although they will be displaced diagonally from where they would be if their efficiencies were normal. However, if the efficiency of one of the reactions deviates more than the other from the mean values, a negative sample may be off from the 60:40 line. For the seven samples characterized by the method invented here (Figure 4, Figure 5) the measured CT-values can be corrected for the differences between their specific PCR efficiencies and the mean efficiencies (eq. 23):

$$CT_{corr} = CT_{meas} * \frac{\log(1 + E)}{\log(1 + \langle E \rangle)}$$

The corrected CT-values are shown with open symbols and they are connected to the measured CT-values by arrows (Figure 6). Although some arrows are diagonal, indicating that the two reactions are inhibited to about the same degree, which does not affect classification, there are some important exceptions.

To account for experimental error and variations in PCR efficiencies in classification of samples, we estimate limits within which negative samples should be found. Keeping the intercept fixed in eq. 20, gives (eq. 24):

$$5 \quad CT_{IgL\kappa} = \frac{\log(1 + E_{IgL\lambda})}{\log(1 + E_{IgL\kappa})} * CT_{IgL\lambda} + \frac{\log(K_{RS} * \frac{N_{0IgL\lambda}}{N_{0IgL\kappa}})}{\langle \log(1 + E_{IgL\kappa}) \rangle}$$

we calculate the standard deviation of the slope,  $k = \log(1 + E_{IgL\lambda}) / \log(1 + E_{IgL\kappa})$ , from the efficiencies determined for the six samples (BR17 was excluded) characterized by *in situ* calibration. This gave  $SD = 0.031$ . For a normal distribution 95% confidence interval is given by  $\text{mean} \pm 1.96 * SD$ . In Figure 3 the dashed lines indicate the interval (eq. 25):

$$10 \quad CT_{IgL\kappa} = (0.946 \pm 0.060) * CT_{IgL\lambda} + 0.021$$

Although the confidence interval takes into account most of the experimental variation, it accounts neither for the variance in the intercept nor the natural variation in the  $IgL\kappa : IgL\lambda$  expression ratio among healthy individuals. These factors would broaden the confidence  
15 interval further. Hence, the interval indicates where negative samples are expected to be found with at least 95 % probability. All negative samples in this study fall within this interval (Figure 3).

Positive samples with  $IgL\kappa$  clonality are below the 60:40 line, while those with  $IgL\lambda$   
20 clonality are above it. Most positive samples fall outside the confidence interval. However, there are some important exceptions. The most striking is BR17, which uncorrected falls within the confidence interval and would be classified as normal. However, after correction for its anomalous PCR efficiencies by the method invented here it falls far outside the confidence interval and can safely be classified as lymphoma with  $IgL\lambda$  clonality (Figure 6  
25 and 7). The reason sample BR5 is within the interval was not established; most likely it is also due to anomalous PCR efficiencies. Sample BR23 has very high CT values, indicating very few copies of both  $IgL\kappa$  and  $IgL\lambda$  cDNA, and was found by IHC analysis to be a T-cell lymphoma.

Example 4. Determination of bcr-abl transcription relative to transcription of GAPDH for CML diagnosis in patient samples using Taqman based real-time PCR assay

Peripheral blood samples from CML patients and controls were extracted at Sahlgrenska University hospital in Gothenburg, Sweden. White blood cells were counted and 100 000 cells were lysed in EL-buffer (Qiagen) and PBS, and stored at -20 until mRNA extraction. RNA-extraction was performed on the Genovision GenoM Robotic Workstation. PolydT coated magnetic beads were used to extract mRNA from lysed blood cells by applying a magnetic force separating the mRNA from other components. The other components are washed away and the mRNA can be eluted by heat. cDNA was synthesized in solution containing 1x Gibco buffer x5, 100mM DDT, 1mM dNTP, 20µM random hexamers, 1 U/µl Rnase inhibitor, 10U/µl Superscript II (Invitrogen). RNase free water was added to a final volume of 50 µl to which 50 µl of mRNA from the extraction step was added. The resulting solution was run in a thermocycler at room temperature for 10min, 42°C for 50min, 70°C for 15min, 95° for 5 min.

Primers used in the BCR-ABL reaction were GCATTCCGCTGACCATCAATA (b2a2-s), TCCAACGAGCGGCTTCAC (b2a2-as) and CCACTGGATTAGCAGAGTTCAA (b3a2-s). The sequence specific probe used was FAM-CAGCGGCCAGTAGCATCTGCTTTGA-BHQ1

Primers used in the GAPDH reaction CAACTGGGACGACTGGAGA (GAPDH-s) and GAAGATGGTGATGGGATTTC (GAPDH-as) and FAM-CAAGCTTCCCGTTCTCAGCC-DQ or FAM-CAAGCTTCCCGTTCTCAGCC-BHQ1 was used as sequence specific probe.

Solutions containing 1x Platinum PCR Buffer (Invitrogen), 4mM MgCl<sub>2</sub> 0.5mM dNTP, 1.25 U Platinum Taq polymerase (Invitrogen), 0.833 µM b2a2-s primer, 0.833 µM b3a2-s primer, 0.833 µM b2a2-as primer, 0.833 µM BCR-ABL probe, and 5µl template from reverse transcription to a total volume of 20µl for the BCR-ABL reaction. The corresponding solution for the GAPDH reaction contained 1x Platinum PCR Buffer (Invitrogen), 4mM MgCl<sub>2</sub> 0.5mM dNTP, 1.25 U Platinum Taq polymerase (Invitrogen),

0.833  $\mu$ M GAPDH-s primer, 0.833  $\mu$ M GAPDH-as primer, 0.833  $\mu$ M GAPDH probe, and 5  $\mu$ l template from reverse transcription to a total volume of 20  $\mu$ l.

5 Samples were run in the Rotorgene (Corbett Research) with fluorescence excitation at 470nm and emission at 510nm. Thermal cycling was programmed at 2min initial denaturation at 95°C and 50-55 cycles of 95°C for 30s and 60°C for 60s.

PCR efficiencies were determined by serially diluting the samples in four steps à two times (Figure 8) for five patient samples (Figure 9).

10

Example 5. Determination of bcr-abl and GAPDH transcription using dye assay  
PCR-product template was prepared by amplification of BCR-ABL and GAPDH fragments in cDNA from K562 cells. The PCR-product was purified using the QIAquick PCR purification kit (Qiagen).

15

Primers used in the BCR-ABL reaction were GCATTCCGCTGACCATCAATA (b2a2-s), TCCAACGAGCGGCTTCAC (b2a2-as) and CCACTGGATTAGCAGAGTTCAA (b3a2-s).

20 Primers used in the GAPDH reaction were CAACTGGGACGACTGGAGA (GAPDH-s) and GAAGATGGTGATGGGATTTTC (GAPDH-as).

Solutions containing 1x Platinum PCR Buffer (Invitrogen), 4mM MgCl<sub>2</sub> 0.5mM dNTP, 1.25 U Platinum Taq polymerase (Invitrogen), 0.833  $\mu$ M b2a2-s primer, 0.833  $\mu$ M b3a2-s primer, 1:80 000 dilution of SYBR Green I, and 6.25  $\mu$ l template from reverse transcription  
25 to a total volume of 25  $\mu$ l for the BCR-ABL reaction (Figure 10). The corresponding solution for the GAPDH reaction contained 1x Platinum PCR Buffer (Invitrogen), 4mM MgCl<sub>2</sub> 0.5mM dNTP, 1.25 U Platinum Taq polymerase (Invitrogen), 0.833  $\mu$ M GAPDH-s primer, 0.833  $\mu$ M GAPDH-as primer, 1:80 000 dilution of SYBR Green I, and 6.25  $\mu$ l template from reverse transcription to a total volume of 25  $\mu$ l (Figure 11).

30

Samples were run in the iCycler (Bio-Rad) with fluorescence excitation at 490nm and detection at 530nm. Thermal cycling was programmed at 2min initial denaturation at 95°C

and 50 cycles of 95°C for 20s, 60°C for 20s, 73°C for 20s. A melt curve was performed from 65°C to 95°C.

# SEQUENCE LISTING

5

SEQ. ID. NO.1

Strand: Single

Nucleic acid

10 PCR primer

5'-TCT CGT AGT CTG CTT TGC TCA - 3'

15

SEQ. ID. NO.2

Strand: Single

Nucleic acid

PCR primer

20

5'-CT CAT CTT TCA CCT CAC CCC - 3',

25

SEQ. ID. NO.3

Strand: Single

Nucleic acid

PCR primer

30

5'- C TCA GGC GTC AGG CTC - 3'

35 SEQ. ID. NO.4

Strand: Single

Nucleic acid

PCR primer

40

5'-C TGC ACT CAA TAA ACC CTC AAT -3'



**CLAIMS**

1. A method for determining efficiency of a PCR wherein the number of copies of a particular nucleic acid sequence in a test sample is determined, comprising amplification of DNA by polymerase chain reaction of the sample itself, or a diluted stock solution of the sample itself, and one or more controlled dilutions of said sample, and registering the number of amplification cycles required to obtain a certain amount of product (CT), and estimating the efficiency of the PCR in the sample from the dependence of CT on the dilution factor.
2. A method according to claim 1, wherein the amounts of two nucleic acid sequences in a sample is compared by determining the PCR efficiencies of the two reactions according to claim 1.
3. A method according to claims 1-2, wherein the ratio of two nucleic acids in a test sample is determined using the relation:

$$\frac{N_{0A}}{N_{0B}} \propto \frac{(1 + \langle E_A \rangle)^{CT_A}}{(1 + \langle E_B \rangle)^{CT_B}}$$

where the CT values are measured in the test sample and the PCR efficiencies  $\langle E \rangle$  are determined separately for a training set of representative samples comprising said nucleic acid sequence by the procedure in claim 1 or an equivalent procedure such as kinetic PCR.

4. A method according to claims 1-2, wherein the ratio of two nucleic acid sequences is determined in a sample using the relation:

$$\frac{N_{0A}}{N_{0B}} = K_{RS} \frac{(1 + E_A)^{CT_A}}{(1 + E_B)^{CT_B}}$$

also taking into account the relative sensitivity of the two PCR assays.

5. A method according to claims 1-2, wherein the amount of a nucleic acid sequence is determined in a biological sample according to either of the claims 1-4; wherein the

nucleic acid is RNA, preferably one or more mRNAs that have been converted to DNA by reverse transcription or a similar process.

6. A method for diagnosing and/or classifying a disease by comparing the expression ratio of two genes by determining the ratio of the corresponding mRNAs in a sample according to either of the claims 1-5.
7. A method according to claim 6, wherein lymphoma is diagnosed by comparing the expression of at least two genes according to either of the claims 1-5, wherein the relative expression of the genes is different in clonal samples compared to healthy tissue.
8. A method according to claim 7, wherein either of the two genes is expressed in each clone of lymphocytes, and are present in a particular ratio in healthy individuals, which ratio is altered in positive samples due to clonality indicating presence of lymphoma.
9. A method according to claim 8, wherein at least a pair of the genes, the expression of which is compared, are the immunoglobulin kappa and lambda light chains.
10. A method according to claim 9, wherein the expression of the immunoglobulin kappa and lambda light chains is compared by determining the  $Ig\kappa : Ig\lambda$  mRNA ratio by reverse transcription PCR, preferably real-time PCR.
11. A method according to claim 6-10, wherein one or more of PCR primers are used that are sufficiently complementary to  
5'-TCT CGT AGT CTG CTT TGC TCA - 3' (SEQ. ID. NO.1), and  
5'-CT CAT CTT TCA CCT CAC CCC - 3' (SEQ. ID. NO. 2), and  
5'- C TCA GGC GTC AGG CTC - 3' (SEQ. ID. NO. 3) and  
5'-C TGC ACT CAA TAA ACC CTC AAT -3' (SEQ. ID. NO. 4), respectively.
12. A method according to claim 11, wherein the degree of complementarity is at least 80%.

13. A method according to claim 1-6, wherein CML is diagnosed by determining the expression of bcr-abl fusion transcript.
- 5 14. A method according to claim 6, wherein the expression of three or more genes are compared.
15. A method for monitoring a disease progress, wherein the expression of two or more genes are compared.
- 10 16. A method for making disease prognosis, wherein the expression of two or more genes are compared
- 15 17. A method for comparing the presence of splicing variants of a gene by determining their relative amounts according to either of the claims 1-5.
18. A method for comparing the activities of alternative promoters by determining the relative amounts of their transcripts according to either of the claims 1-5.
- 20 19. A method for determining the amount of virus or bacteria in a sample according to either of the claims 1-5.
20. Method for diagnostic testing for cancer, including lymphoma, wherein at least the kappa:lambda expression is determined.



1/11

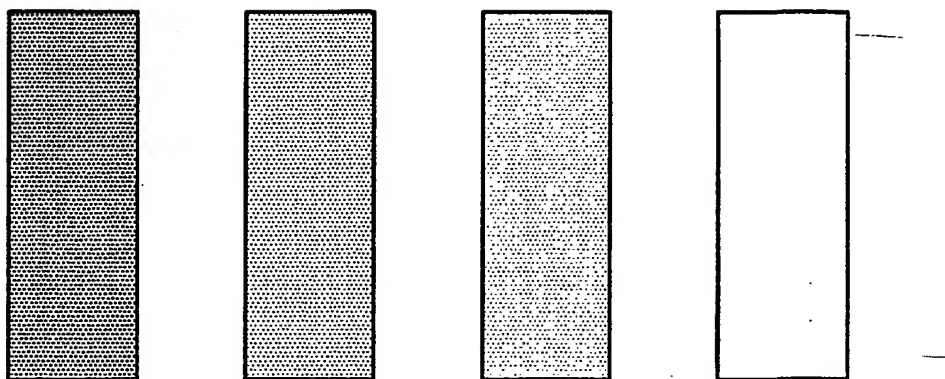


FIG. 1

Patient sample nr.	Efficiency BCR-ABL	Efficiency GAPDH
1	87 —	74
2	78	61
3	71	66
4	73	55
5	67	55

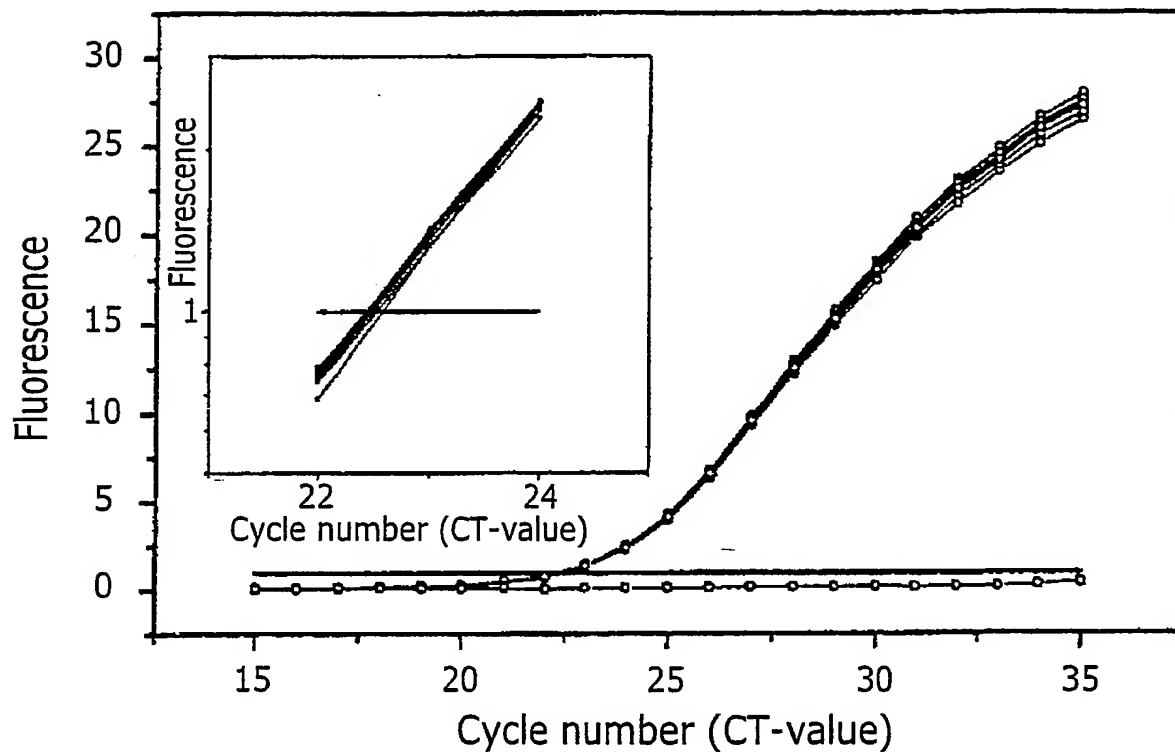
FIG. 9



2/11

IgL $\kappa$  intra-assay

A

IgL $\lambda$  intra-assay

B

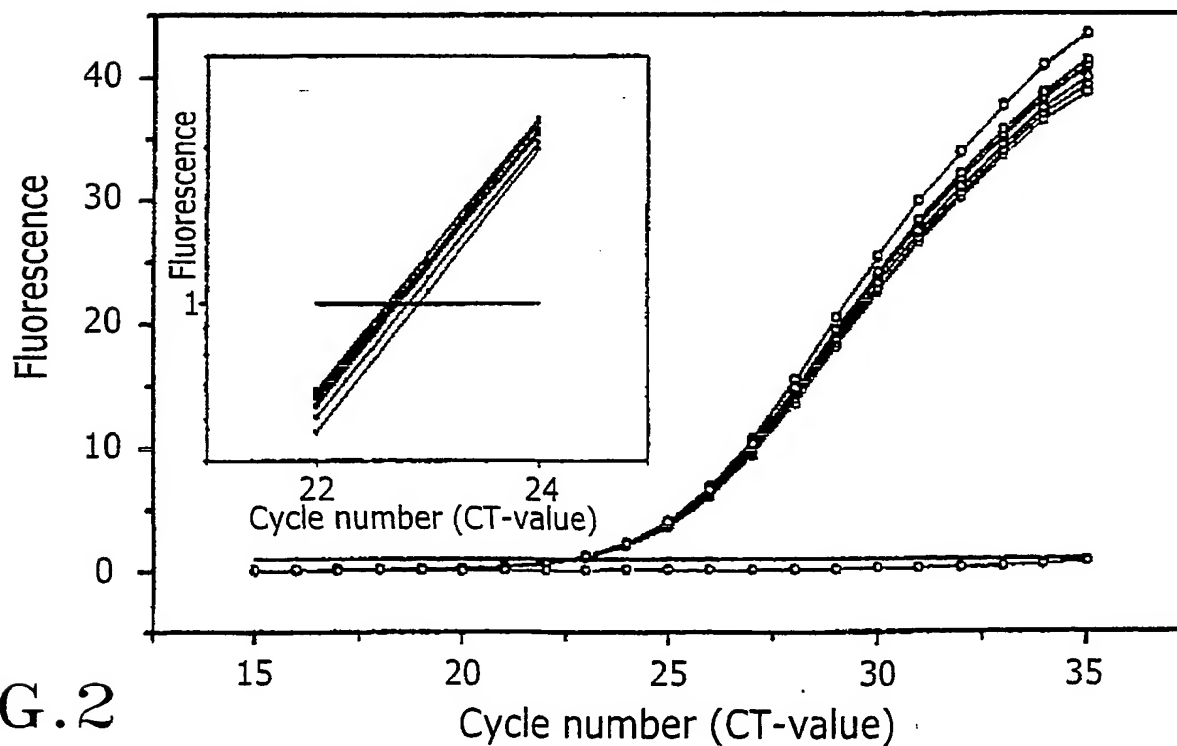


FIG.2

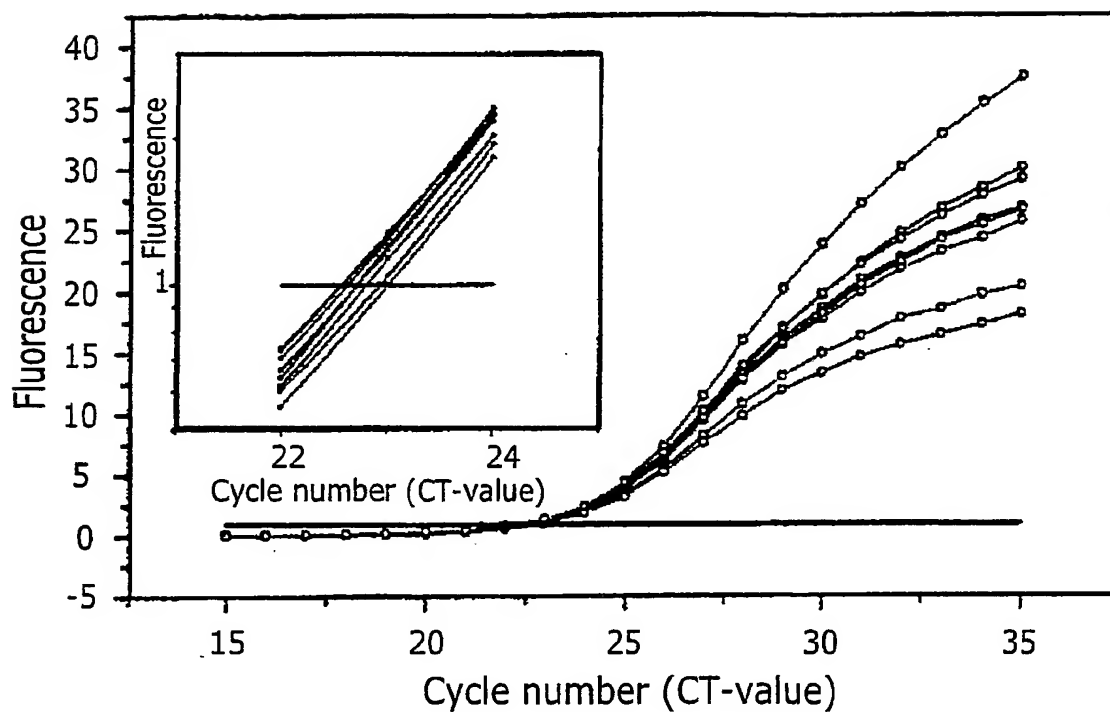




3/11

IgL $\kappa$  intra-assay

C

IgL $\lambda$  intra-assay

D

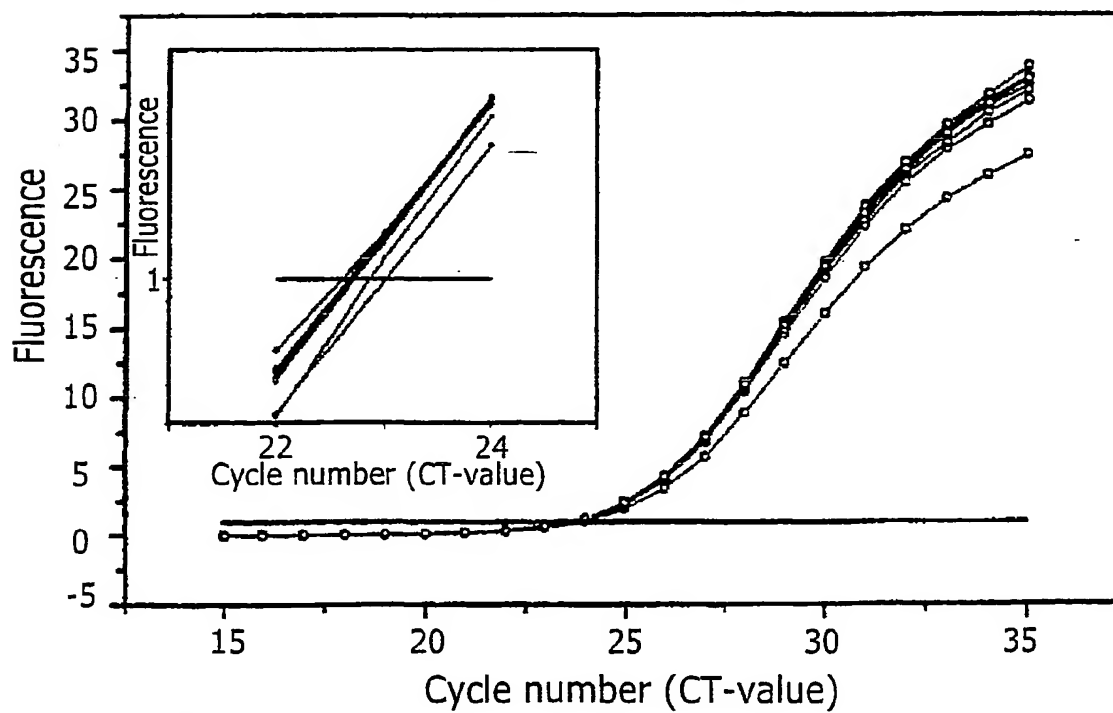


FIG.2 cont.



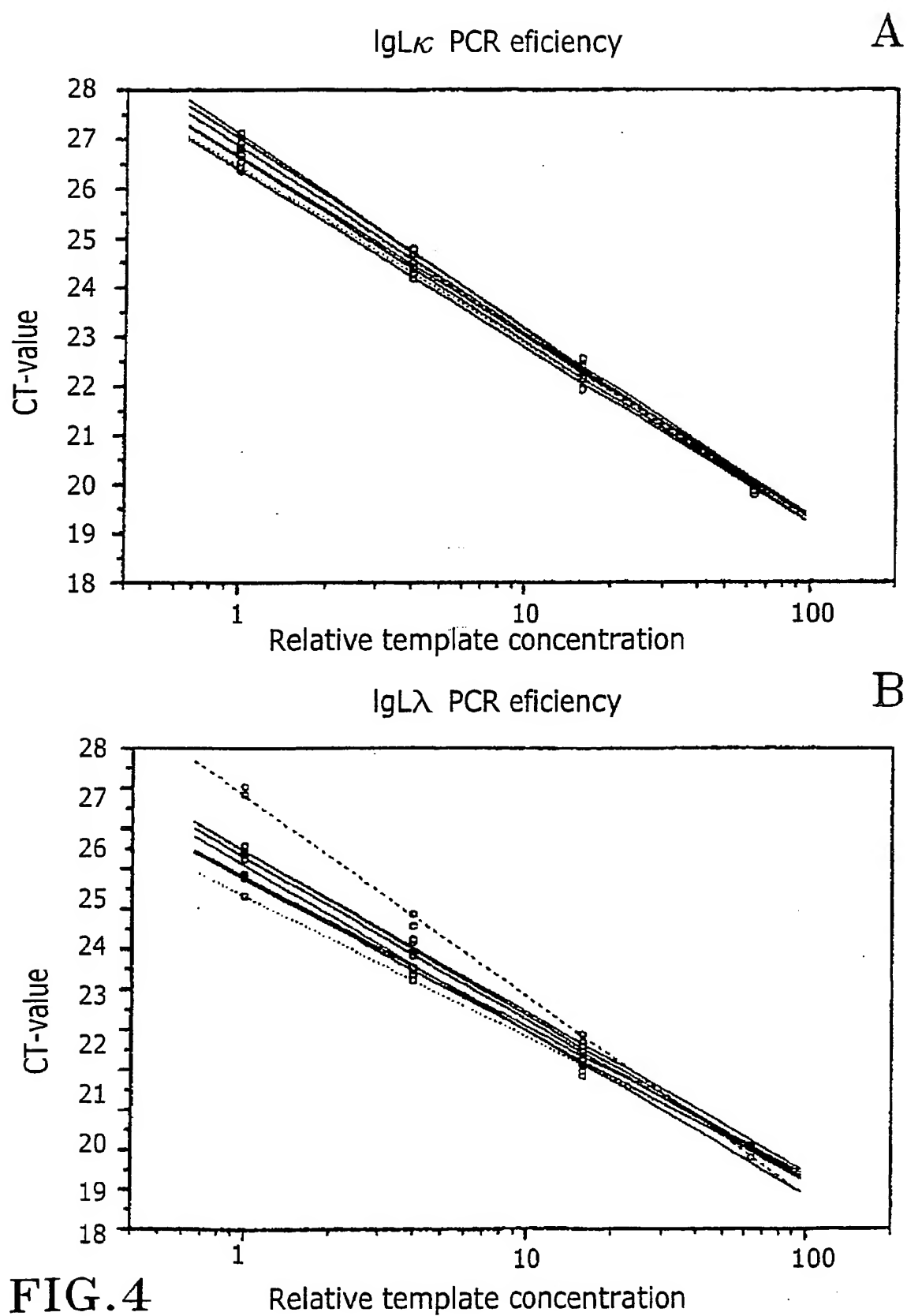
4/11

	IgL $\kappa$			IgL $\lambda$		
	CT <sub>max</sub> <sup>-</sup> CT <sub>min</sub> <sup>*</sup>	CV(CT) <sup>+</sup>	CV(copy) <sup>+</sup>	CT <sub>max</sub> <sup>-</sup> CT <sub>min</sub> <sup>*</sup>	CV(CT) <sup>+</sup>	CV(copy) <sup>+</sup>
Intra-assay	0,15	0,23%	3,0%	0,28	0,39%	4,9%
Intra-assay	0,37	0,62%	8,1%	0,25	0,37%	5,0%

FIG.3



5/11





6/11

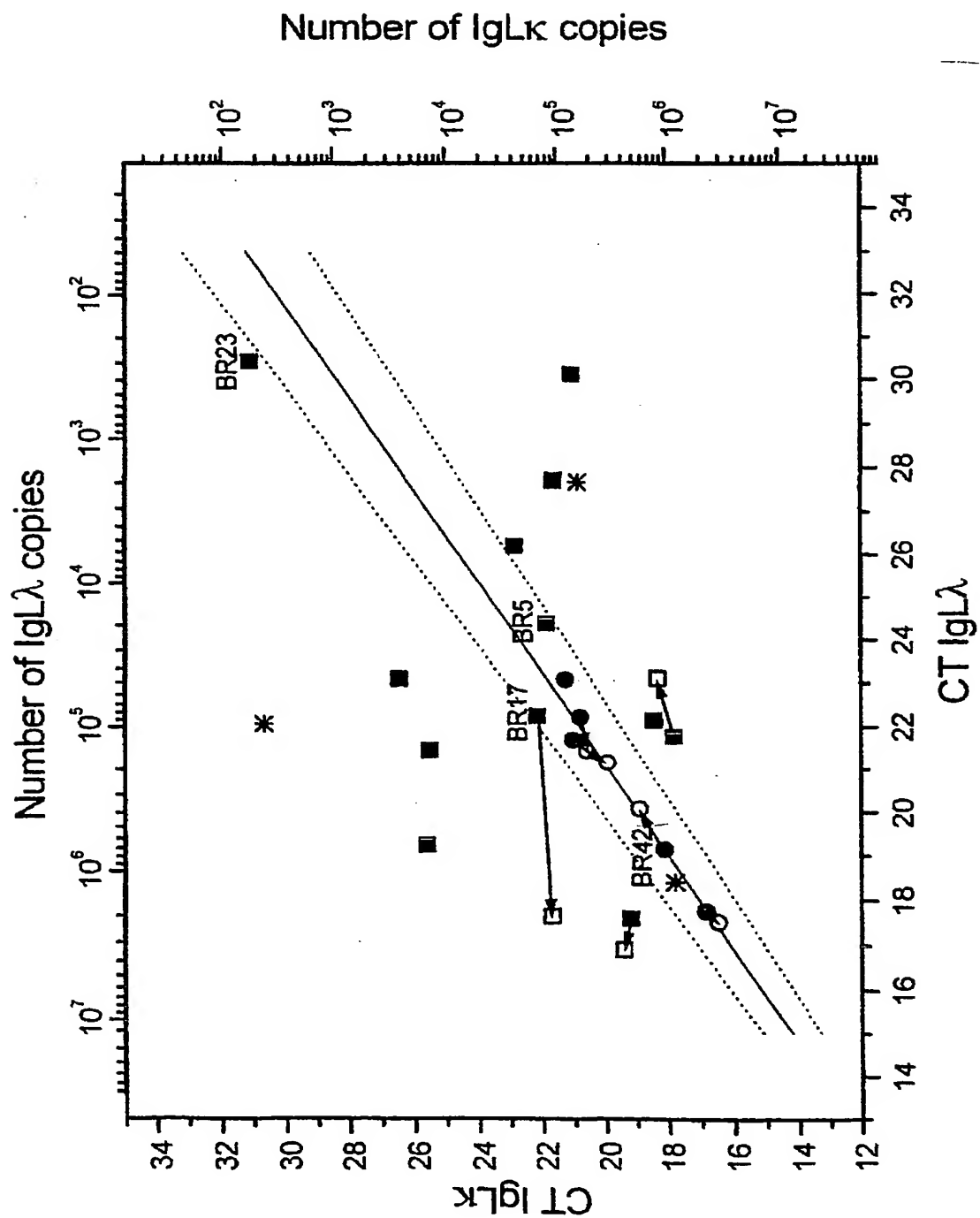
Patient sample	IgLκ		IgLλ		$\frac{1 + E_{\text{IgL}\kappa}}{1 + E_{\text{IgL}\lambda}}$	K <sub>RS</sub> <sup>†</sup>
	CT*	E <sup>†</sup>	CT*	E <sup>†</sup>		
BR0	21.0	79.4	21.7	76.8	1.014	1.42
BR18	17.2	86.6	18.8	75.2	1.065	-
BR37	22.2	84.2	20.8	75.7	1.048	1.84
BR42	18.7	90.4	17.9	84.4	1.033	1.41
BR46	17.1	82.8	16.4	77.9	1.028	1.42
BR52	21.9	88.7	17.9	85.8	1.015	-
Mean	85.4	79.3			1.034	1.52
Template	94.7	93.2			1.008	-
BR17	23.1	83.0	23.3	58.9	1.152	-

FIG.5





7/11





8/11

Sample	Real-time PCR % $\kappa$ expression <sup>†</sup>	R.E.A.L. classification <sup>*</sup>	IHC clonality <sup>*</sup>	Flow Cytometry $\kappa$ : $\lambda$ expression <sup>†</sup>
BR0	62 (52)	Lymphadenitis	poly	ND
BR1	(6)	NHL-CLL	IgL $\lambda$	1:99
BR2	(0.4)	NHL-DLBCL	IgL $\lambda$	3:97
BR4	(99)	NHL-CLL	IgL $\kappa$	95:5
BR5	(76)	NHL-Foll	IgL $\kappa$	65:35
BR6	(98)	NHL-DLBCL	IgL $\kappa$	ND
BR8	(8)	NHL-Foll	IgL $\lambda$	ND
BR11	(2)	NHL-Foll	IgL $\lambda$	32:68
BR13	(96)	NHL-CLL	IgL $\kappa$	96:4
BR17	6 (43)	NHL-Foll	IgL $\lambda$	32:68
BR18	16 (24)	NHL-Foll	IgL $\lambda$	ND
BR22	(83)	NHL-Foll	IgL $\kappa$	ND
BR23	(40)	T-cell lymphoma	-	ND
BR37	56 (63)	Lymphadenitis	poly	ND
BR42	62 (60)	Lymphadenitis	poly	ND
BR43	(68)	Lymphadenitis	poly	ND
BR46	62 (59)	Lymphadenitis	poly	46:54
BR48	(88)	NHL-Foll	IgL $\kappa$	ND
BR52	93 (89)	NHL-CLL	IgL $\kappa$	95:5

FIG. 7



9/11

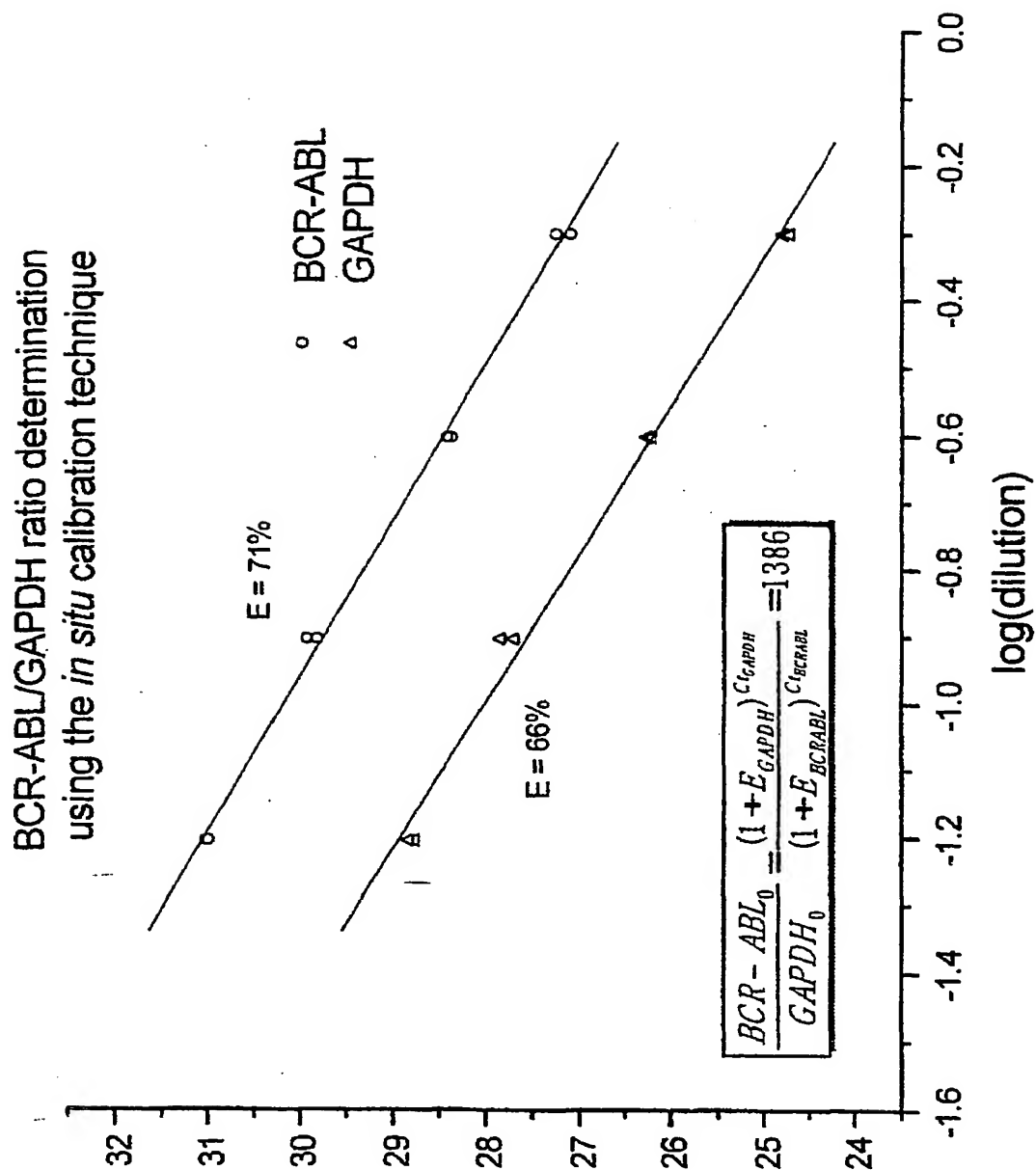


FIG.8



10/11

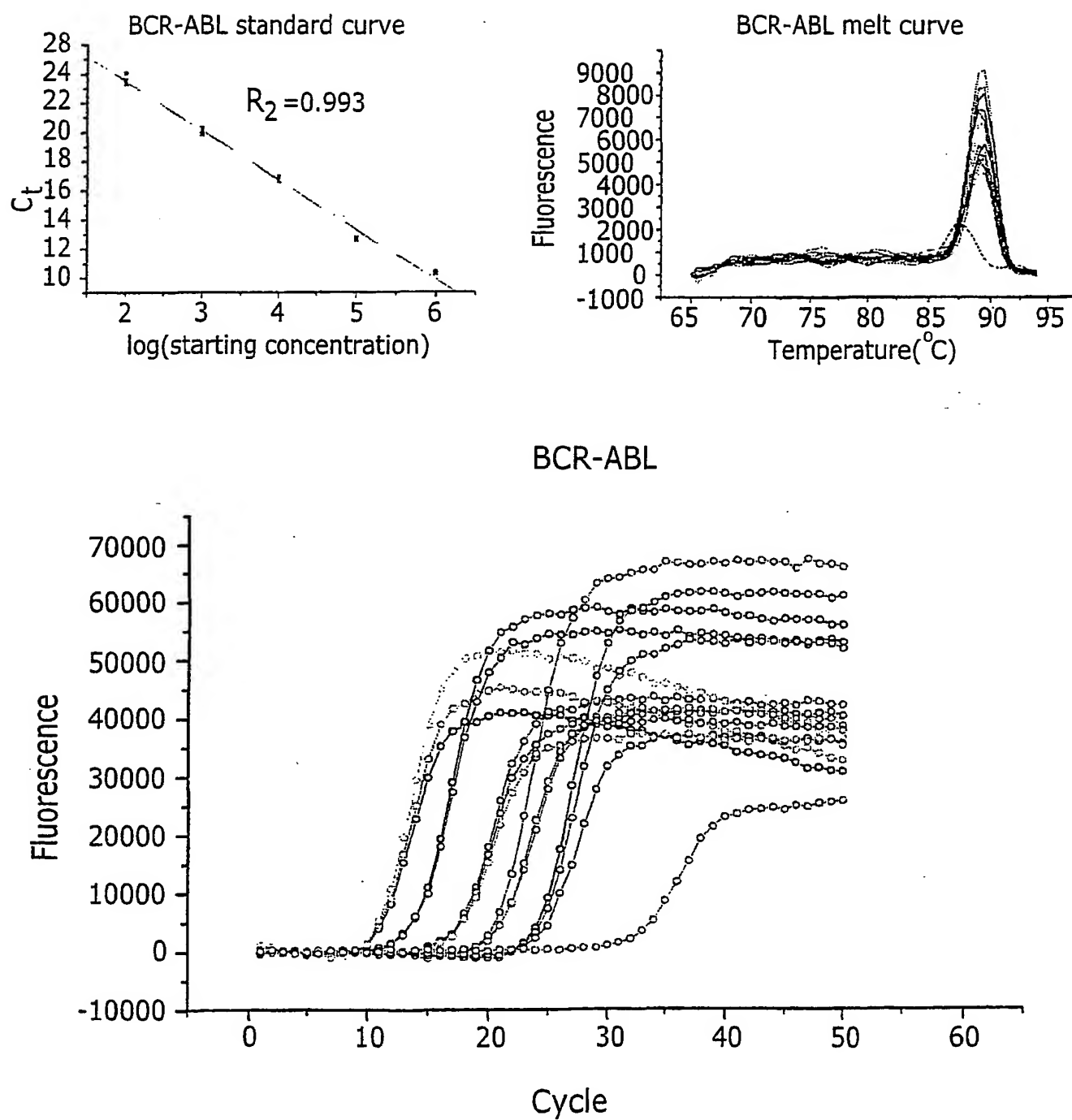


FIG. 10





11/11

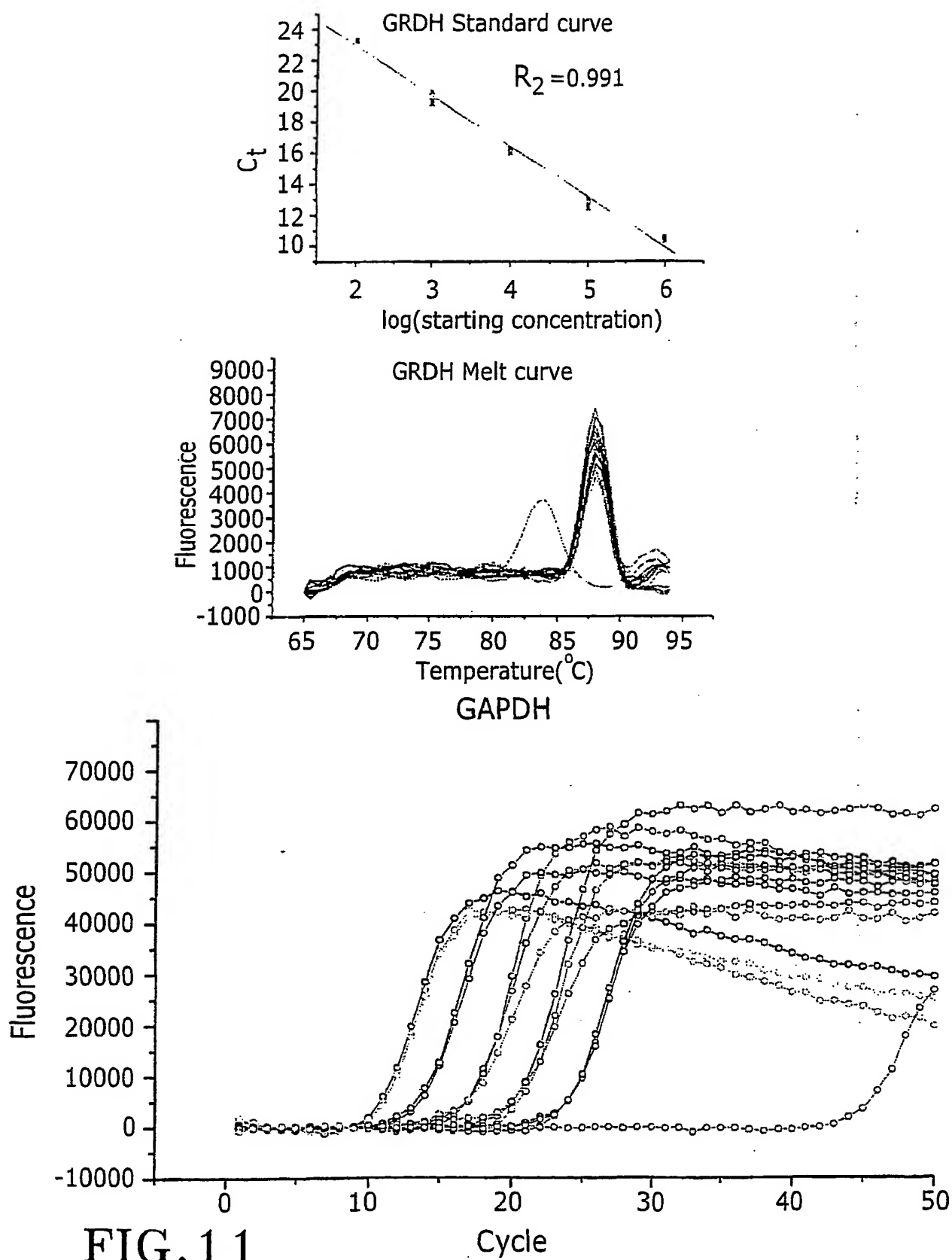


FIG. 11



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/SE 02/01093

## A. CLASSIFICATION OF SUBJECT MATTER

IPC7: C12Q 1/68 // C07K 16/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7: C12Q, C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-INTERNAL, WPI DATYA, MEDLINE, BIOSIS

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 0044935 A2 (BAVARIAN NORDIC RESEARCH INSTITUTE A/S), 3 August 2000 (03.08.00), page 20, line 15 - line 23	1-6,13-19
Y	--	7-12,20
Y	Modern Pathology, Volume 7, No 4, 1994, (USA), Yao-Tseng Chen et al., "Clonality Analysis of B-cell Lymphoma in Fresh-frozen and Paraffin-embedded Tissues: The Effects of Variable Polymerase Chain Reaction Parameters", abstract	7-12,20
	--	

☒ Further documents are listed in the continuation of Box C.☒ See patent family annex.

\* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier application or patent but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

Date of mailing of the international search report

20 Sept 2002

26 -09- 2002

Name and mailing address of the ISA/  
Swedish Patent Office  
Box 5055, S-102 42 STOCKHOLM  
Facsimile No. +46 8 666 02 86

Authorized officer

Frida Plym Forshell/LR  
Telephone No. +46 8 782 25 00

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/SE 02/01093

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>SAMOSZUK, M.K. ET AL.: Limitations of Numerical Ratios for Defining Monoclonality of Immunoglobulin Light Chains in B-Cell Lymphomas. Diagnostic Immunology 3:133-138 (1985). See abstract</p> <p>-----</p>	9-12,20

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/SE02/01093**Box I** Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2. ☒ Claims Nos.: 15, 16, 20  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:  
  
see extra sheet
  
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II** Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see extra sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☒ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/SE02/01093

## Continuation of Box I.2

Present claims 15-16 and 20 relate to an extremely large number of possible methods. Support within the meaning of Article 6 PCT and disclosure within the meaning of Article 5 PCT is to be found, however, for only a very small proportion of the methods claimed. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible.

Consequently, the search has been carried out for those parts of the claims which appear to be supported and disclosed, namely methods where the expression of two or more genes are compared using the equation suggested in claim 3.

## Continuation of Box II.2

In the present application, claim 1 is the only claim that relates to a method for determining the efficiency of a PCR. All the other claims relate to methods for determining or comparing amounts of nucleic acids. Thus, none of the claims 2-20 can be subordinate to claim 1 and therefore, claim 2 has the status of an independent claim. Furthermore, claims 5, 13 and 19 relate to the determination of the amount of a single nucleic acid. These claims can not be subordinate to the method for comparing amounts presented in claims 2-4. Therefore, claim 5 also has the status of an independent claim.

According to Article 34 (3) (a-c) and Rule 13.2, an international application shall relate to one invention only or to a group of inventions linked by one or more of the same or corresponding "special technical features", i.e. features that define a contribution which each of the inventions makes over the prior art.

Three different inventions have been identified, namely:

1. Claim 1 relating to a method for determining the efficiency of a PCR.
2. Claims 2-4, 6-12, 15-18 and 20 relating to methods where amounts of nucleic acids are compared.
3. Claims 5, 13 and 19 relating to methods for determining the amount of nucleic acid in a sample.

These inventions lack a common technical feature that defines a contribution over the prior art.

## International application No.

PCT/SE 02/01093

Form PCT/ISA/210 (patent family annex) (July 1998)

